

# Adaptive Comparative Judgment (ACJ)

## Presenters

### Mary Power

[mary.power@uwaterloo.ca](mailto:mary.power@uwaterloo.ca)

<https://ca.linkedin.com/in/mary-power-b796449>

### Tonya Elliott

[tonya.elliott@uwaterloo.ca](mailto:tonya.elliott@uwaterloo.ca)

<https://ca.linkedin.com/in/tonelliott>

Note: will be on parental leave May 11 until June 2018

## Overview and Key Journal Articles

Decent overview c/o Wikipedia: [https://en.wikipedia.org/wiki/Adaptive\\_comparative\\_judgement](https://en.wikipedia.org/wiki/Adaptive_comparative_judgement)

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4), 273-286.

- Article in which CJ was first referenced
- No such thing as an “equal” judgment
- Brought back into modern-day use by Pollitt & Murray in 1993  
Pollitt, A, & Murray, NJ (1993) What raters really pay attention to. Language Testing Research Colloquium, Cambridge. Republished in Milanovic, M & Saville, N (Eds), *Studies in Language Testing 3: Performance Testing, Cognition and Assessment*, Cambridge: Cambridge University Press.

Pollitt, A. (2012). The method of adaptive comparative judgement. *Assessment in Education: Principles, Policy & Practice*, 19(3), 281-300.

- Seems to be the “go to” article for this topic (as it pertains to teacher scoring). Relies on experts’ implicit knowledge of “better”. Designed for better validity (holistic grading), although also touted as being beneficial when inter-marker reliability is low.
- “CJ may be used as a test equating function as well as a test scoring function”
- Could also be used to narrow down/shortlist a bunch of entries (e.g., for an award/competition)

Jones, I., & Wheadon, C. (2015). Peer assessment using comparative and absolute judgement. *Studies in Educational Evaluation*, 47, 93-101.

- Interesting results re: reliability and validity of comparative vs. absolute peer assessments.
- Peer assessment can be used for one of five reasons (social control tool; assessment tool; learning tool; ‘learn-how-to-assess-tool’; active participation tool). This paper focuses on peer assessment as an assessment tool.
- Encouraged students to talk about their judgments because previous studies indicated that judging was an enjoyable task, they were keen to talk about their judgments, and the researchers believed that talking would help foster a positive learning environment
- Page 23 (emphasis added): “One possible fruitful direction for comparative judgement is peer assessment as a learning tool, engaging students in discussion and consideration of what constitutes a high-quality response to an open-ended or ambiguous assessment task. There is cautious optimism for believing that comparative approaches to learning may be more beneficial than absolute approaches (Pachur & Olsson, 2012). However, one possible limitation to this approach is that comparative judgement does not necessarily provide descriptive feedback to students. Comparative judgement studies have provided peer assessors with the opportunity to provide a text-based comment when making each judgement (Seery et al., 2012). However, in our experience that **slows the judgement process and produces comments that lack insight.**”

Jones, I., & Alcock, L. (2012). Summative peer assessment of undergraduate calculus using adaptive comparative judgement. *Mapping university mathematics assessment practices*, 63-74.

- Some surprising results re: expert vs. novice vs. peer reliability & validity results. Lots of interesting comments and survey responses.
- “A rank order produced by ACJ can be used to allocate grades to students in the standard way. This can be done using norm referencing, for example, allocating the top 20% of scripts a grade ‘A’ and so on. Alternatively, it can be done using criterion referencing. This requires sampling scripts from across the rank order and comparing them against agreed assessment criteria. Boundary scripts within the rank order can then be identified and grades applied accordingly. In our case the students will eventually be awarded grades using criterion referencing, but that process was not within the scope of this project.”
- Interesting result re: looking at expert (math PhD) vs. novices (arts PhD) as controls: “The relatively strong correlation between the novice and two other groups leads to the counter-intuitive and unexpected conclusion that novices lacking knowledge of advanced mathematics can, to some extent at least, assess understanding of advanced mathematics”
- Some students were interviewed and commented that that learning benefits did occur. “We note that students’ responses to the final question, ‘How did you find the experience overall?’, suggest that they found judging peers’ scripts challenging, but beneficial for learning.”

## Tools and Other Pilots

**Digital Assess:** <http://digitalassess.com/>

The Digital Assess platform was the original adaptive comparative judgement system and Dr. Alastair Pollitt (from above articles) is on their advisory board. It’s sold commercially and used to be called “tag assessments” or TAG. Reached out to them several times (web forms, emails, LinkedIn), but never received a reply

**No More Marking:** <https://www.nomoremarking.com/>

This tool focuses more on the reliability/validity/efficiency-related benefits of ACJ (vs. the learning-related benefits – i.e., it does include functionality to provide students with feedback). You can, however, sign up for free on their website and follow their 5-minute guide for a demo: <http://knowledge.nomoremarking.com/getting-started/the-5-minute-guide>. Dr. Chris Wheadon is the founder and Dr. Ian Jones is the sr. scientific advisor (from above articles). Very responsive to emails.

**ComPAIR (UBC):** <http://ubc.github.io/compair/site/index.html>

This is the tool we’re investigating for our pilot (very early days, but code is open source):  
<https://github.com/ubc/compair>

Learning-related research in progress @UBC: <https://sclt.science.ubc.ca/supperseries/adaptive-comparative-judgment-100-year-old-novel-take-peer-assessment>

## Additional Readings

Jones, I., Swan, M., & Pollitt, A. (2015). Assessing mathematical problem solving using comparative judgement. *International Journal of Science and Mathematics Education*, 13(1), 151-177.

Jones, I., & Inglis, M. (2015). The problem of assessing problem solving: can comparative judgement help?. *Educational Studies in Mathematics*, 89(3), 337-355.

McMahon, S., & Jones, I. (2015). A comparative judgement approach to teacher assessment. *Assessment in Education: Principles, Policy & Practice*, 22(3), 368-389.

Newhouse, C. P. (2014). Using digital representations of practical production work for summative assessment. *Assessment in Education: Principles, Policy & Practice*, 21(2), 205-220.

Pollitt, A., & Wingfield, M. (2012). Adaptive comparative judgement: Adapting adaptive assessment to assess the quality of students’ work. *International Journal of e-Assessment*, 2(2).

Whitehouse, C. (2012). Testing the validity of judgements about geography essays using the adaptive comparative judgement method.